Title	点過程・空間パタンの最尤法に於ける非正則漸近性の例について(統計的漸近理論とその応用)
Author(s)	尾形, 良彦
Citation	数理解析研究所講究録 (1983), 507: 33-43
Issue Date	1983-12
URL	http://hdl.handle.net/2433/103761
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
507巻 1983年 33-43

点過程・空間パタンの最尤法に於ける非正則漸近性の例について

統計数理研究所　　尾形良彦
Yosihiko Ogata

§1 統計力学における Hard-core モデルの最尤法

平面中の領域 V で定義された pairwise ポテンシャル $\phi(r)$ に対する Gibbs canonical 分布は

$$f(x_1,\cdots,x_N) = \exp(-\sum_{i<j}\sum \phi(|x_i - x_j|)/Z_N,\ x_i \in V,\ i=1,2,\cdots,N.$$

で与えられる。但し Z_N は normalizing constant で

$$Z_N = \int_{V^N} \exp(-\sum_{i<j}\sum \phi(|x_i - x_j|)dx_1\cdots dx_N \qquad (1)$$

である。いまポテンシャル ϕ_σ がパラメタ σ によって特徴付けられたとすると対数尤度は

$$\log L(\sigma) = -\sum_{i<j}\sum \phi_\sigma(|x_i - x_j|) - \log Z_N(\sigma) \qquad (2)$$

Hard-core potential は

$$\phi_\sigma(r) = \begin{cases} 0 & r \geq \sigma \\ \infty & r < \sigma \end{cases} \qquad (3)$$

で定義されたものである。Hard-core は最も簡単なモデルのひとつでありその漸近理論についての文献も少なくない。たと

えば Gates & Westcott (1980), Moran (1979), Ripley & Silverman (1978), Saunders & Funk (1979), Westcott (1982) などがある。

Ogata and Tanemura (1983) によると Hard-core ポテンシャルの場合，次の性質を満たす正の単調増加なψ-関数が存在する：すべての $\tau=N\sigma^2/V$ に対して

$$\frac{1}{N}\log\left\{Z_N(\sigma)/V^N\right\}=-\int_0^{\tau}\psi(t)/t\,dt。 \qquad (4)$$

従って定数項を除いて，ポアソンパタンに対する，尤度比は次のようになる。

$$\log L(\sigma)=\begin{cases}-\infty & \sigma>\min\{|x_i-x_j|;\ i\neq j\}\\ N\int_0^{\tau}\psi(t)/t\,dt & \sigma\leqq\min\{|x_i-x_j|;\ i\neq j\}\end{cases} \qquad (5)$$

これから，ただちにわかるように Hard-core model の最尤推定量は $\hat{\sigma}_N=\min\{|x_i-x_j|;\,i\neq j\}$ である。この漸近分布を知りたい。また真のパラメタが σ_0 としたときの尤度比統計量

$$(-2)\log\hat{\lambda}=(-2)\left\{\log L(\sigma_0)-\log L(\hat{\sigma}_N)\right\} \qquad (6)$$

の漸近的な性質が知りたい。

いま $\bar{Z}_N(\sigma)=Z_N(\sigma)/V^N$ とおく。(6)を書き直すと(3),(4)に

よって

$$
\begin{aligned}
(-2)\log\hat{\lambda} &= (-2)\left\{\log Z_N(\hat{\sigma}_N) - \log Z_N(\sigma_0)\right\} \\
&= 2N\int_{\rho\sigma_0^2}^{\rho\hat{\sigma}^2} \psi(t)/t\, dt
\end{aligned}
\tag{7}
$$

となる。但し $\rho = N/V$ である。ρ を一定のまま $N\to\infty$（または $V\to\infty$）のとき $\hat{\sigma}_N \to \sigma_0$（in prob.）であることに注意する。以下２通りの帰無仮説を考える。

(i) 帰無仮説が Hard-core（すなわち $\sigma_0 > 0$）のとき。

充分大きな N に対して (7) から次のような漸近式が期待できる。

$$
(-2)\log\hat{\lambda} \sim 2N\,(\hat{\sigma}_N^2 - \sigma_0^2)\cdot\psi(\rho\sigma_0^2)/\sigma_0^2 \tag{8}
$$

次のような Survivor 関数を考える。

$$
\begin{aligned}
S(x;\sigma_0) &= P_{\sigma_0}\left\{N(\hat{\sigma}_N - \sigma_0) > x \,\middle|\, \text{hard-core, 直径 } \sigma_0\right\} \\
&= P_0\left\{|x_i - x_j| \geq \sigma_0 + \frac{x}{N} \text{ for all } i,j\,(i\neq j)\right\} \Big/ P_0\left\{|x_i - x_j| \geq \sigma_0 \text{ for all } i,j\,(i\neq j)\right\} \\
&= Z_N\left(\sigma_0 + \frac{x}{N}\right) \Big/ Z_N(\sigma_0)。
\end{aligned}
\tag{9}
$$

ここで P_0 はポアソン確率測度である。例えば下から２番目の式の分母を解釈すると「V の中に N 個の点をデタラメ（ポア

ソン）に放り投げたとき，N個の点のどの2つともσ_0以上の距離を保つ確率」のことであり，これを式であらわすと$\Xi_N(\sigma_0)=Z_N(\sigma_0)/V^N$ で与えられる。何故なら $Z_N(\sigma_0)=Volume\{(x_1,\cdots,x_N)\in V^N; |x_i-x_j|\geqq\sigma_0,\ i\neq j\}$。

$\rho=N/V$を一定として$N, V\uparrow\infty$とすると(4), (9)式によって漸近的に

$$\log S(x;\sigma_0)=-N\int_{\rho\sigma_0^2}^{\rho(\sigma_0+x/N)^2}\psi(t)/t\,dt\ \sim -\{2\psi(\rho\sigma_0^2)/\sigma_0\}\cdot x \tag{10}$$

このことから$N(\hat{\sigma}-\sigma_0)$は漸近的に平均値が $\sigma_0/\{2\psi(\rho\sigma_0^2)\}$の指数分布になることがわかる。また(8), (9), (10)によって

$$P_{\sigma_0}\{(-2)\log\hat{\lambda}<x\}=1-e^{-x/2} \tag{11}$$

であり，これは自由度2のカイ二乗分布（平均値2の指数分布）である。ここで関係したパラメタが一個（σのみ）であることに注意すれば通常の尤度比統計量の振舞は自由度1のカイ二乗であることが期待されたが，そうでないことがわかった。

(ii) 帰無仮説が Poisson（すなわち$\sigma_0=0$）のとき。

(7)から始めるが，(8)のかわりに次の関係式

$$(-2)\log\hat{\lambda}=2N\int_0^{\rho\hat{\sigma}_N^2}\psi(\tau)/\tau\,d\tau=2N\cdot\frac{\pi}{2}\cdot\rho\,\hat{\sigma}_N^2 \tag{12}$$

が導びかれる。ここに$\frac{\pi}{2}$は hard-core model の第2次ヴィリアル係数であることに注意する。同様にして(9)をつかって(10)のかわりに

$$\log S_0(x) \sim -\frac{\pi}{2}\frac{\rho}{N}x^2 = -\frac{\pi}{2}\frac{1}{V}x^2 \tag{13}$$

が $N\uparrow\infty$ の漸近的に期待できる。ただしVは領域の面積である。もしたとえばVを半径1の球面というように固定すると$V=4\pi$ だから $\log S_0(x) \sim -x^2/8$ が得られこれは Moran (1979) の結果の一部である。(13) は言いかえると

$$P_0(N\hat{\sigma}_N < x) \sim 1 - e^{-(\pi/2V)x^2} \tag{14}$$

である。これをつかって (12) の分布を調べると

$$P_0\{\log\hat{\lambda} < y\} = P_0\left\{\frac{\pi}{V}(N\hat{\sigma}_N)^2 < y\right\} = P_0\left\{N\hat{\sigma}_N < \sqrt{\frac{V}{\pi}y}\right\} \sim 1 - e^{-y/2} \tag{15}$$

となり前と同じ分布が導びかれた。

§2 自己制御点過程モデルの最む法

以上はパラメタの境界でも内点と同様む度比統計量の漸近分布が，非正則とは言え，一致する例であった。以下は一次元の点過程でパラメタの境界と内点では振舞が全く異なる話である。条件付危険度関数 (conditional intensity function, 以

下 CIF と書く）が

$$\lambda(t) = \exp\left\{\alpha + \beta(t - \rho N(t)\right\} \qquad (16)$$

で特徴付けられる一次元点過程を考える。ここで $N(t)$ は時間区間 $[0, t]$ で発生した点の数であり，α, β, ρ は推定されるべきパラメタである。このCIFの対数は，発生点の間では傾き β で線形に増大し発生点のところで γ だけ落下する。このモデルは Isham & Westcott (1979) の自己制御 (self-correcting) モデルの特殊な場合にあたり，Vere-Jones (1978) が提案した応力解放モデルとは，落下サイズが確率変数であることを除けば，同様のものである。

CIF が存在すれば常に対数尤度が存在して

$$\log L = \int_0^T \log \lambda(t)\, dN(t) - \int_0^T \lambda(t)\, dt \qquad (17)$$

のように書ける。$[0, T]$ は観測時間区間。最尤法の正則漸近理論は Ogata (1978) や Kutoyants (1979, 1980) に議論されている。(16) はすぐわかるように $N_T > 0$ ならば必ず唯一の最大尤度が存在する。この漸近的な性質を調べるために以下の2つの場合を考慮せねばならない。

(i) $\beta_0 > 0$ のとき

(16) を次のようにパラメタ変換を行なうと便利である。

$$\log \lambda(t) = \alpha + \beta X(t) + \gamma t/T \tag{18}$$

但し $X(t)=t-N(t)$ で真のパラメタは $\theta=(\alpha_0, \beta_0, 0)$ となるように時間のスケールをも調節した。$X(t)$ はマルコフ過程であり，その骨格（skelton）$X(n)$ は格子上に値をとる。適当な $0 \leq \nu < 1$ について $Z(t)=X(t-\nu)$ によって $X(n)$ が整数値をとるマルコフ連鎖とすることができる。そこで Tweedie（1982a, b）の定常分布がモーメントをもつための判定条件をつかうと $Z(n)$ はすべてのモーメントを持つ $\left[E_\pi\{|Z(n)|^\sigma \cdot \exp \rho |Z(n)|\} < \infty \right]$ ことがわかり，$Z(t)$, $X(t)$ も同様の性質をもつ。また $Z(n)$ は irreducible の aperiodic でエルゴード的である。詳細は Vere-Jones & Ogata（1983）を参照。

以上の準備のもとに確率積分の重み付き大数の弱法則を示し，また Kutoyants（1979）のマルチンゲール中心極限定理の為の条件を点検することによって次のような正則な結果を得る。

<u>Prop. 1</u>　（18）のパラメタの最尤推定量（$\sqrt{T}(\hat{\alpha}_T - \alpha_0)$, $\sqrt{T}(\hat{\beta}_T - \beta_0)$, $\sqrt{T}\hat{\gamma}_T$）は漸近的に正規で分散共分散行列 J^{-1} に従う。ただし

$$J = \begin{pmatrix} U & V & U/2 \\ * & W & V/2 \\ * & * & U/3 \end{pmatrix}, \quad \begin{aligned} U &= E_\pi\left[\int_0^1 e^{-\beta Z(t)}dt\right], \\ V &= E_\pi\left[\int_0^1 Z(t)e^{-\beta Z(t)}dt\right], \\ W &= E_\pi\left[\int_0^1 Z(t)^2 e^{-\beta Z(t)}dt\right]。\end{aligned} \tag{19}$$

証明については Ogata & Vere-Jones (1983) 参照。

(ii) <u>$\beta_0 = 0$ のとき</u>

このときは (16) を (18) とも違って次のようなパラメタ変換をすると見易い。

$$\log \lambda(t) = \alpha + \beta X(t)/\sqrt{T} + \gamma t/T \tag{20}$$

ここで $X(t) = t - N(t)$ は上と同じだが帰無仮説は $\theta_0 = (0, 0, 0)$ としても一般性を失なわない。このとき $N(t)$ はポアソン過程だから，$t = \tau T$ とおいて

$$W_T(\tau) = \left\{N(\tau \cdot T) - \tau \cdot T\right\}/\sqrt{T}, \quad 0 < \tau < 1 \tag{21}$$

を定義すると $T \uparrow \infty$ に従って $W_T(\tau)$ は Wiener 過程 $W(\tau)$, $0 < \tau < 1$ に弱収束することがわかっている。すなわち任意の $D[0,1]$ 上の連続な汎関数 $F(\cdot)$ に対して $F(W_T) \to F(W)$。このことから Fisher-score と Hessian 行列について次のような漸近分布が得られる。

<u>Prop. 2</u>　帰無仮説 $\theta_0=(1,0,0)$ のもとで

$$\frac{1}{\sqrt{T}}\frac{\partial \log L}{\partial \theta_0} \Rightarrow \begin{pmatrix} W(1) \\ -\frac{1}{2}\{W(1)^2-1\} \\ W(1)-\int_0^1 W(\tau)d\tau \end{pmatrix},\quad \frac{1}{T}\frac{\partial^2 \log L}{\partial\theta_0\partial\theta_0'} \Rightarrow \begin{pmatrix} 1 & \int_0^1 W(\tau)d\tau & \frac{1}{2} \\ * & \int_0^1 W^2(\tau)d\tau & \int_0^1 \tau W(\tau)d\tau \\ * & * & \frac{1}{3} \end{pmatrix}$$

(22)

ここで注目すべき異常性は Fisher-score の第2成分の周辺分布が正規でないことと Hessian 行列の 6 つの成分が定数にではなくて分布に収束していることである。これにもかかわらずその他の最尤法の性質は通常どおりである。すなわち，極限 Hessian 行列（Fisher 行列）は正定値性（positive-definiteness），一致性（consistency），対数尤度の Taylor 展開の 3 次以上の項が 0 に確率収束することなどが成立する。たとえば，Fisher 行列の 2 次形式は，$\xi=(\xi_1,\xi_2,\xi_3)$ に対して，(22) から

$$\xi \Xi \xi' = \int_0^1 (\xi_1+\xi_2 W(\tau)+\xi_3\tau)^2 d\tau > 0 \qquad (a.e.) \quad (23)$$

と書けるので正定値行列である。詳細は Ogata & Vere-Jones (1983) を参照。

最後にポアソンを帰無仮説として次のようなモデルによって尤度比検定する時の漸近分布について議論してみる

$$\lambda(t) = \rho \exp[\beta(t - \rho N_t)], \tag{24}$$

但しρの推定値として$\hat{\rho} = N(T)/T$を代入するものとする。いま記号$W_T^0(\tau) \equiv W_T(\tau) - \tau W_T(1)$を定義すると上の議論と同様にして

$$2\log(L/L_1) \sim \left[4\int_0^1 W_0(\tau)^2 d\tau\right]^{-1} \tag{25}$$

であることが導びかれる。ここに$W_0(\tau) \equiv W(\tau) - \tau W(1)$は Brownian bridge と呼ばれている。Anderson & Darling (1952) に Cramér-von Mises 統計量の適合度検定のための

$$\omega^2 = \int_0^1 W_0(\tau)^2 d\tau \tag{26}$$

の数値表が与えられているので利用できる。

References for Section 1

Gates, D. J. and Westcott, M. (1980). Further bounds for the distribution of minimum interpoint distance on a sphere. Biometrika, 67, 466-469.

Moran, P. A. P. (1979). The closest pair of N random point on the surface of a sphere. Biometrika, 66, 158-62.

Ripley, B. D. and Silverman, B. W. (1978). Quick tests for spatial interaction. Biometrika, 65, 3, 641-642.

Saunders, R. and Funk, G. M. (1977). Poisson limits for a clustering model of Strauss. J. Appl. Prob. 14, 776-784.

Westcott, M. (1982). Approximations to hard-core models and their application to statistical analysis. 65th Birthday Volume for Professor P. A. Moran (J. Appl. Prob. Supplement, No. 1, 281-292.).

References for Section 2

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. Ann. Math. Statist. 23, 193-212.

Isham, V. and Westcott, M. (1979). A self-correcting point process. Stoch. Proc. Appl. 9, 335-347.

Kutoyants, Yu. A. (1979). Local asymptotic normality for processes of Poisson type. Izvest. Akad. Arm. Nauk. Ser. Matekatika, 14, 3-20.

——(1980). Estimation of Parameters of Stochastic Processes (in Russian), (Aremenian Akad. Nauk.).

Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimates for stationary point processes. Ann. Inst. Stat. Math. 30, 243-261.

Ogata, Y. and Vere-Jones, D. (1983). Inference for earthquake models: - a self-correcting model, preprint.

Tweedie, R. L. (1982). Criteria for rates of convergence of Markov chains, with application to queueing and storage theory.(to appear J. Appl. Probab.).

——(1982). The existence of moments for stationary Markov chains. (submitted to J. Appl. Probab.).

Vere-Jones, D. (1978). Earthquake prediction - a statistician's view. J. Phys. Earth, 26, 129-146.

Vere-Jones, D. and Ogata, Y. (1983). On the moments of a self-correcting process, preprint.